ED 373 754 IR 016 764

AUTHOR Ross, Steven M.; Casey, Jason
TITLE Using Interactive Software To Develop Students' Problem-Solving Skills: Evaluation of the "Intelligent Physics Tutor."
PUB DATE 94
NOTE 13p.; In: Proceedings of Selected Research and Development Presentations at the 1994 National Convention of the Association for Educational Communications and Technology Sponsored by the Research and Theory Division (16th, Nashville, TN, February 16-20, 1994); see IR 016 784.
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Academic Achievement; *Computer Assisted Instruction; Computer Softw... Evaluation; *Courseware; Formative Evaluation; High Schools; Instructional Innovation; *Intelligent Tutoring Systems; Physics; *Problem Solving; *Secondary School Science; Student Reaction; Teacher Response
IDENTIFIERS Interactive Systems

ABSTRACT

This study presents a formative evaluation of "The Intelligent Physics Tutor," an interactive software program designed to improve high school physics students' problem-solving skills. The purposes of this evaluation were to describe how the "Tutor" was used in the pilot high school class; to determine teacher and student reactions to the program in terms of weaknesses, strengths, and suggestions for improvement; and to examine possible impacts of the "Tutor" on problem-solving strategy and performance. The test group comprised of three honors physics classes taught by the same teacher; only one of the three classes used the "Tutor." Multiple data sources used in the evaluation included teacher experiences and attitudes and student attitudes, achievement, and problem-solving strategies. These data were obtained via surveys, interviews, and teacher observation. Findings from the semester-long tryout indicated that both the teacher and the students were positive about the "Tutor's" instructional benefits. The most favorable reactions were given by lower-ability students who had trouble solving problems on their own. While there were no differences in achievement between the "Tutor" class and two control classes, the "Tutor" students demonstrated relatively greater use of systematic problem-solving strategies over time. (Contains 5 references.) (JLB)

Title:

Using Interactive Software To Develop Students' Problem-Solving Skills: Evaluation of the *Intelligent Physics Tutor*

Authors:

Steven M. Ross
Jason Casey

Memphis State University
MEMPHIS, TN 38152

Bowen Loftin
University of Houston-Downtown

Abstract
The present study was a formative evaluation of the *The Intelligent Physics Tutor*, an interactive software program designed to improve physics students' problem-solving skills. Multiple data sources used in the evaluation included teacher experiences and attitudes; and student attitudes, achievement, and problem-solving strategies. Findings from a semester-long tryout in a high school physics class indicated that both the teacher and the students were very positive about the *Tutor*'s instructional benefits. The most favorable reactions were by lower-ability students who had trouble solving problems on their own. While there were no differences in achievement between the *Tutor* class and two control classes, *Tutor* students demonstrated relatively greater use of systematic problem-solving strategies over time.

Introduction

As orientations in instructional theory and practice continue to evolve from their behavioral roots, cognitive and "constructivist" viewpoints have had increasing influence on the design of computer-based instruction (CBI) (Bednar, Cunningham, Duffy, & Perry, 1991; Di Vesta & Rieber, 1987). Such viewpoints emphasize the processes by which students learn higher-order skills and the attitudes and interests they develop through their experiences The marriage today of cognitive perspectives with interactive, automated delivery systems orients today's designers toward considering how learner needs and mental models can be diagnosed on task and instruction adapted accordingly. The present project, involving the design and evaluation of an intelligent tutoring system for high school physics, was grounded on this orientation.

The Intelligent Physics Tutor Project

The Intelligent Physics Tutor project (Loftin, 1990), was sponsored by NASA/Johnson Space Center and Apple Classrooms of Tomorrow. Its purpose is to create an interactive environment, using the architecture of an intelligent tutoring system, in which physics students can practice problem-solving applications while receiving individualized guidance and feedback. Consequently, as students work on the Tutor in class, the teacher is freed to give special attention to individuals who need it. The *Tutor* operates on an Apple Macintosh II computer and makes use of high resolution graphics and sound.

The basic instructional model is designed to diagnose students' misconceptions while they solve problems. The identified misconceptions then form the basis for determining the immediate feedback provided on the current problem and for selecting additional problems to be worked. The analysis takes place via a "local strategy" that records every response (as "student action codes") that the student makes in solving the problem. Using the work space, an available calculator, the keyboard, and given variables and equations that can be selected for the problem solution, the student attempts to work the problem. At each step, a comparison is made between the student's action and the correct or expected action. If a discrepancy is indicated, the most likely misconception for that problem is diagnosed on the basis of the error response and the cumulative action codes. Appropriate feedback, as outlined below, is then provided.

Students receive on-task feedback messages only when an error is detected; otherwise they are allowed to proceed without interruption. An example of an error feedback display is shown in Figure 1. The instructor pictured is Dr. Bowen Loftin, the *Tutor* developer; individual teachers can scan their own pictures into the program to personalize this feedback.

The error messages provided are geared to the specific misconception diagnosed and become increasingly specific as errors reflecting that misconception accumulate. Importantly, by correcting the error at the time it is made, the system prevents the learner from proceeding down an incorrect pathway (and thus losing time and experiencing frustration). Based on the cumulative frequencies of misconceptions recorded for the individual across completed problems, the *Tutor*'s "global model" then selects a new problem that requires practicing skills directly related to the diagnosed areas of weaknesses (i.e., the specific misconceptions having the highest frequencies of occurrences).

Additional feedback systems provide summaries and analyses of performances for both classes (see Figure 2) and individuals (see Figure 3). The basic feedback display shows the

misconceptions currently diagnosed while suggesting a prescription of how many related exercises (or problems) should be given for strengthening the problem-solving skills for a given misconception. Parallel feedback for the class as a whole provides useful guidance for whole-group followup instruction.

Evaluation Objectives

The purposes of the present evaluation were threefold. One was to describe how the *Tutor* was being used in the pilot high school class. This information was considered important for (a) identifying possible strategies for integrating the *Tutor* with conventional instructional methods, and (b) understanding the application context in which results occurred. A second purpose was to determine student and teacher reactions to the *Tutor* with regard to weaknesses, strengths, and suggestions for improvement. Based on this information, more confident decisions can be made about needs for "fine-tuning" the user interface and program content. A third purpose was to examine possible impacts of using the *Tutor* on problem-solving strategy and performance. This interest included both the *processes* (how a problem was solved) and *products* (accuracy of solutions) of problem-solving activities.

In interpreting the results, readers should be aware of several limitations of the present evaluation context. First, the population tested consisted of honors physics students attending a middle-class suburban high school. These students were more capable and experienced problem solvers than are typical, beginning physics students at most schools in the United States. Second, as will be described below, the inability to randomly assign students to "treatment" and control classes necessitated the use of a "quasi-experimental" design, i.e., comparison of pre-existing classes. Third, only one teacher was involved, thus confounding *Tutor* outcomes with her personality, orientation, and style. Given that these factors, the present evaluation was designed to be primarily "formative" in nature, in the sense of identifying how the Tutor operates and its "likely" effects. A summative evaluation (testing its "effectiveness") would be the next logical step in the research and development process.

Evaluation Methodology

Participants and Design

A quasi-experimental design consisting of three intact classes of students was employed. All were honors physics classes taught by Ms. Beverly Lee at Clear Creek High School. Unfortunately, due to chance factors, the three classes differed somewhat in size and in ability. It was therefore decided to provide a strong rather than weak test of the *Physics Tutor* though the selection of the lowest-ability (and second largest) class as the PT group. This class was taught sixth period and consisted of 19 students. The PT group worked problems on the Tutor following regular instruction in four course units (one-dimensional motion, two-dimensional motion, one-dimensional dynamics, and dynamics). In most instances, students worked on Tutor problems initially in pairs and later independently. The rationale for this procedure was both practical, to increase practice opportunities on the available computers; and pedagogical, to give students the opportunity to help each other during the beginning stages of problem-solving.

The "conventional problems" control group ($n = 13$) was established in the first period class, which was the smallest in size and the highest in ability. This group received conventional textbook and teacher-made problems on the selected units.

The second-period class consisted of 25 students. It received print versions of the same *Physics Tutor* problems administered to the PT group and, like the PT group, worked on the problems

initially in pairs and later independently. However, these students were not afforded the advantages of immediate feedback, diagnosis, and adaptive problem selection available to the PT group. This group will be referred to as the "no-feedback" group. Both this group and the control group had limited computer activity through their work with a simulation program early in the year. Neither group had any exposure to the *Tutor*.

Demographic and pre-course achievement data were available for describing and comparing the three classes on (a) the mathematics subtest of the Stanford Achievement Test taken the previous spring, (b) a knowledge pretest administered in the physics class at the beginning of the year, and (c) the highest-level mathematics course completed. Inferential statistical analysis (one-way analysis of variance) was used for comparing classes on these pre-course variables. Although the PT and no-feedback class scored directionally lower than the control class on the two tests, no significant differences were indicated.

Instrumentation

Two times during the year (September and November), a reaction survey was administered to students in class. Part 1 of the survey, which was administered to all three classes, consisted of six questions dealing with feelings about physics and problem solving. Examples are "I like learning physics" and "I feel confident about my ability to solve physics problems." Responses were made on a five-point Likert scale, ranging from "strongly disagree" to "strongly agree."

Two additional sections of the survey were administered to PT students only. Part II contained six ratings items dealing specifically with experiences in using the *Tutor*; e.g., "The procedures for using the *Tutor* were easy to follow" and "I prefer the *Tutor* problems to doing textbook problems." Part III consisted of two open-ended questions asking what was liked best and least about using the *Tutor* for the current lesson.

An overall survey was administered to the PT class in February, several weeks after they had completed their last Tutor exercises. This survey consisted of 24 ratings items dealing with specific features of the Tutor (e.g., understandability, operation of the mouse and calculator, problem selection, graphics).

In addition to the survey, approximately five randomly selected students from each of the classes were interviewed in September and November. For the two control groups, the interview consisted of four questions concerning attitudes toward physics as a subject, solving physics problems, and their physics course. Students in the PT group were asked those questions along with five additional questions concerning experiences in using the *Tutor*: activities, likes and dislikes, impact on problem solving, and suggestions for improvement.

During the fall semester, the teacher kept a journal in which she recorded her impressions and activities in using the Tutor in the sixth-period class. Information about the curriculum, instructional objectives, and student characteristics was also given.

Students' semester test results, as scored by the teacher, were recorded for analysis purposes. The specific tests were 1st six-weeks test, 2nd six-weeks test, 3rd six-weeks test, semester exam, and final average. The scores analyzed were total scores (0-100%) without any breakdowns for subsections or specific problems. Also, to explore whether students in the three classes differed in their approaches to physics problem solving, a scheme was devised by the class teacher (Beverly Lee) and the Tutor developer (B. Loftin) for evaluating how

individual problems were set up and worked. The system involved identifying specific steps or features of a correct, systematic problem-solving procedure for the particular problem types administered (e.g., statement of correct formula, identification of variables, etc.). Discussion and informal tryout of an initial model led to the development of the final scheme consisting of 25 features (e.g., initial position, velocity, final position, etc.). Two physics teachers from local high schools, who were trained to use the scheme, demonstrated a very high rate of agreement (Overall $M = 94\%$) in their scoring.

The same two judges evaluated 28 problems that were selected from different tests on the basis of having varied degrees of structural similarity with actual *Tutor* problems. Specifically, some were exact repetitions of Tutor problems, and some were similar to Tutor problems but contained different contexts. Analyses examined results on individual problems and on clusters of problems representing similar problem types (e.g., one-dimensional kinematics, two-dimensional dynamics, etc.).

Results

Student Surveys

Results on the student surveys showed generally positive and consistent reactions toward physics as a subject and the physics class. No significant differences between classes were found on any of the items. The strongest trend for class variations was indicated on Item 2-Round 2, "Confident in ability." Total agreement ("agree" + "strongly agree") on that item was highest for the PT group (94%) compared to the no-feedback (84%) and control (64%) group. Also noteworthy was that the PT and no-feedback classes became somewhat more negative across rounds about working with a partner (from 6% and 4% disagreement, respectively, to 20% and 11% disagreement). In general, the PT group was somewhat more negative toward physics and problem-solving during Round 1 than were the two control classes, but was comparable to the controls by Round 2.

Reactions by the PT group to the Tutor were highly positive on both rounds. In fact, the highest percentage of students expressing disagreement on any item was only 18% which occurred on Round 1 in reacting to "able to work problems" and "prefer *Tutor* to homework." In Round 2 the highest level of disagreement was 13%, on "procedures were easy to follow." In reacting to the statement, "Overall, I liked using the *Tutor*," 92% of the students agreed (only 1 disagreed). Specific features that received the highest percentage of agreement responses were "helped to accurately read problem" (92%), "helped me learn to set up problems" (85%), "liked Yabba Dabba Doo sound" (85%), "initial instructions were sufficient" (84%), and "makes physics fun" (83%). On the other hand, only 53% percent found the calculator "easy to use," 48% agreed that "feedback helps to identify mistakes, "16% thought the "help button offers useful assistance," and 62% agreed that the "screen design tends to clutter as I work." These less positive reactions seem at least partly attributable to students' lack of familiarity with the software. In any case, they indicate relatively mild degrees of dissatisfaction and are clearly outweighed by the clearly positive reactions on nearly all items.

Student Interview

Student interview responses supported the survey findings by showing generally positive attitudes toward the physics course and considerable liking of the Tutor. When asked what they liked best about the Tutor, students most frequently mentioned the immediate feedback. Other responses given by more than one student were "easier to work with formulas," "gives

step-by-step instruction, and the "picture helps to understand the problem." Least liked features identified by more than one student were "too slow," "too many windows, "and "calculator is hard to use."

The last question was whether the *Tutor* helped them to become better problem solvers. Nearly all students strongly agreed. Among the reasons given were demonstrations of the problem, review, and immediate feedback. There was a clear difference between higher- and lower-achievers in the reactions expressed. Students who felt comfortable about their problem-solving abilities felt that the Tutor was useful for initial learning and practice, but slowed them down once they gained proficiency. Specifically, they didn't feel that they needed to work through all the steps that the *Tutor* required. They understood those steps well enough to skip them. As "expert" problem solvers, they were at the stage where they relied on heuristics (general strategies) and did not need step-by-step procedures.

Contrasting views were expressed by four students who described themselves as "struggling" due to weaknesses in mathematics background. They felt that the Tutor was extremely helpful at all stages of the learning process. The main advantage was guiding them through systematic problem-solving procedures, while not allowing them to proceed down an incorrect pathway. Because of these features, they felt that their approaches to problem-solving were changing. Specifically, they were more confident, more reflective and systematic, and less concerned about mathematics than about procedures. All felt that they were able to transfer these new approaches to problem-solving to textbook and examination problems.

Teacher Journal and Interview Reactions

The main advantage of the Tutor, in the teacher's view, was helping students to become more systematic problem solvers by writing down correct formulas and input variables before attempting solutions. For example, on October 21, the teacher wrote:

> One student had not identified the initial horizontal velocity (same magnitude). The program would not allow him to begin working with a formula until he had identified that piece of information, since it would be needed later...The [error] messages he had received all referred to the concept of velocity...Once he had identified that value [V_{ox}], he could use the formula. **This is exactly what the program should be doing!!! The student received the correct messages and had to THINK** [*bold type and capitalization, hers*].

Later (November 8), the teacher discussed how PT students began to accept (though reluctantly at first) the need to complete each step, no matter how small or obvious, in the problem-solving process. Students in the control classes were seemingly less receptive to and less practiced at this orientation.

Not all aspects of the *Tutor* experience were positive, however. Sometimes (November 3) the *Tutor* problems were too few in number or did not operate correctly. Other limitations were excessive time needed for equipment set-ups, equipment failures, and minor problems that students experienced with the interface (getting lost in "too many windows"). The teacher's feeling was that these difficulties were infrequent and minor in importance. She expected their impact to increasingly diminish as students and teachers became more familiar with using the *Tutor* and the few remaining "bugs" in the software were corrected.

In the interview, the teacher expressed the overall feeling that this year's uses of the *Tutor* were much better than last year's. Her hope was that students would be able to transfer what they had learned about problem-solving to paper-and-pencil problems. Although the steps

that the *Tutor* required took extra time, that sacrifice was clearly worthwhile if students became more systematic problem solvers. In her opinion, the high achiever who was successful with both procedures and answers probably did not need the *Tutor* nearly as much as the lower achiever.

Based on her early observations of test results (in November), she felt that some students did not transfer the problem-solving procedure outside the problem domain. With additional practice, however, they seemed to perform better, so there was hope that transfer would increase with experience.

The teacher felt that students looked forward to the days when the computer was used and to their turns on it. Without cooperative learning, it would probably have been more difficult for them to learn the problem-solving skills. They conversed extensively before making responses, especially on new or complex problems.

In characterizing the overall experience, the teacher indicated that "there have been no disappointments." Set-up work and equipment problems were more of an investment at the beginning of the year and with new topics. While she was very pleased with having access to the *Tutor*, she did not feel that she substantively changed her teaching style in employing it. Nor did she believe that the primary value or impact of the *Tutor* lied in engendering such changes. That is, whether or not the Tutor were available, she would have maintained her primary methods of lecture, demonstration, cooperative practice and independent practice. What the Tutor did provide, in her view, was a highly valuable tool for facilitating applications of these strategies through its functions for demonstration, practice, feedback, and record-keeping. While qualitative changes in teaching methods may not occur, the teacher believed that they did take place in student learning, namely, the adoption of more reflective and systematic problem-solving approaches.

Achievement

Class performances were compared on the following achievement tests: 1st six-weeks, 2nd six-weeks, 3rd six-weeks, semester exam, and final average. To determine whether group differences were statistically significant, a one-way ANOVA was conducted on each measure. None of the achievement comparisons was significant, thereby, suggesting that the three classes performed comparably on the five overall achievement measures. Nor were there any significant class differences when test scores were adjusted for pre-course aptitude variables.

Problem-Solving

Problem-solving approaches by students in the three classes were analyzed in several ways. The most narrow and molecular data category consisted of mean scores on individual *elements* across all problems. A second, broader category examined *problem* means by tabulating the number of elements represented in the given problem solution. The third and broadest category combined similar problems representing a certain type of application (e.g., 1D-kinematics, 2D- kinematics, etc.) to yield an overall *problem type* score.

For *problem* scores., results showed an uneven pattern, with significant class differences obtained on only 5 out of the 28 problems and on the overall score (points summed across all problems). On the five significant problems, the PT group's adjusted mean was directionally highest on one, second highest on one, and lowest on three. The no-feedback group was highest on three, second highest on one, and lowest on one. The control group was highest on one,

second highest on three, and lowest on one. On total score, the order of groups was no-feedback ($M = 407$), PT ($M = 394$), and control ($M = 387$).

For *problem-type* scores, only one out of the five comparisons was statistically significant. On the 2-D kinematics problems, the PT group was lowest and the no-feedback group was highest. However, in an effect that approached significance, the PT group was highest on 2D-dynamics-inclined plane problems, whereas the control group was lowest. The direction of the means on the various problem types indicated a trend for the PT group to score relatively higher compared to the other groups as problem complexity (and course experience) increased. Specifically, PT scores were lowest of the three groups on 1-D kinematics problems, but highest on the more difficult 2-D dynamics problems.

For *element* scores, significant effects were obtained on 18 out of 187 elements associated with 1-D kinematics problems (the easiest level). The direction of the means showed the PT group to typically be lowest of the three groups (in 72% of the cases), and the no-feedback group to be highest (in 94%). Results for 2-D kinematics problems showed significant differences on 20 out of 56 elements. As with 1-D kinematics problems, the PT group typically scored lowest (in 95% of the cases) and the no-feedback group highest (in 65%). On 1-D kinematics with 1-D dynamics problems, 9 out of 56 elements showed significant group effects. Here, the results were less consistent, with the PT group scoring highest on 2 elements (22%), second highest on 4 (44%), and lowest on 3 (33%). Finally, significant group differences were obtained on 17 out of the 191 elements on the 2-D dynamics problems: the PT group scored the highest of the three groups on 15 elements (88%). The no-feedback group was second in nearly all cases.

Conclusions

Conclusions based on the results for student attitudes, teacher experiences, and student problem-solving outcomes are described below.

Student Attitudes

Judging from our five independent visits of PT classes during the past two years, and student survey and interview responses, there is no question that students felt very positively about using the Tutor. They greatly looked forward to "Tutor periods" as opportunities to break from the routine of working problems by hand without the benefit of error-sensitive feedback. Working with a computer appeared to be as motivating for these advanced high school students as for the elementary school children we've observed in numerous contexts. There is also no question that the cooperative learning format used by the teacher enhanced the experience for most students, especially lower-achievers who, if forced to work alone, would have struggled with identifying correct problem solving steps. The latter impression is important in considering instructional uses of the *Tutor*, given that the *Tutor* primarily "reacts" to student responses but does not directly "teach" the physics concepts needed to understand the problem. Cooperative learning allows students to discuss (and teach each other) the problem-solving strategies and principles, and then see the results of their decisions.

Negative reactions to the Tutor were infrequently expressed. In fact, not one PT student indicated that he/she preferred conventional methods over the PT problem-solving activities. The dissatisfactions that were conveyed tended to concern minor aspects of the user interface that made the Tutor less convenient to use than paper-and-pencil modes. The most noteworthy of these concerned the ease of using the calculator and the

Help function, the slowness of the system, and the cluttering of the display windows. These factors seemed to become less important as students acquired more experience in using the *Tutor*. In the November interviews, for example, several students characterized the interface procedures as difficult (or new) at first, but "second nature" now.

Our overall conclusion is that no major revisions in the user interface are indicated by the evaluation results. The program developers should review the student reactions, in relation to considerations of the time and cost requirements of associated programming changes, and decide which, if any, revisions would be desirable.

An additional conclusion supported by the student reactions is that the *Tutor* offers greater benefits to lower-achieving than to higher-achieving students. This interpretation is hardly surprising, given that high achievers are, by definition, relatively successful learners who have performed well using traditional learning methods and resources. The literature on problem solving (e.g., Chi, Feltovich, & Glaser, 1982) suggests that expert problem solvers tend to use a "top-down" approach, which emphasizes understanding of problem patterns and contexts, as opposed to a "bottom-up" approach, which emphasizes the piece-by-piece building of the problem solution. From this perspective, it is reasonable that the highest-achieving students sometimes felt constrained by the step-by-step procedures that the *Tutor* enforced. They felt that they were able to skip many of these steps and consequently, solve the problem more efficiently using paper and pencil. All students agreed, however, that beginning a unit with a few *Tutor* problems served a valuable function, by exposing them to a complete, systematic procedure.

It is important to remember that the present subjects were all physics honors students attending a middle- to upper-middle-class suburban school. Thus, the proportion of high-achievers, who experience success working independently, was likely to be substantially greater for this sample than would occur in "regular" physics classes at other schools.

Teacher Attitudes

As conveyed by the teacher journal and interview results, the teacher was very satisfied with the *Tutor* software. She shared the students' perception that the Tutor problems were especially valuable at the beginning of a new unit and for students who were experiencing difficulties. During the year, she encountered few difficulties with the operation of the software or hardware.

Because only one teacher participated in the evaluation, it is risky to make general conclusions about how other teachers would utilize or react to the *Tutor*. Some teachers, for example, might prefer individual problem solving over cooperative activities; some might use the class and individual feedback reports to a greater or lesser extent. Given the ease of using the Tutor and the few problems that were encountered in the present application, our strong impression is that most teachers would react quite positively to the *Tutor* and welcome its availability for their classes.

Unique to the present application, however, was the teacher's long-term involvement with the development of the *Tutor*, interest in the project, and access to the developers who worked locally at NASA. New users, not associated with the project, would be less knowledgeable about and adept at using the program. Support material, such as a "user's manual," is needed to provide guidance on both operating procedures and instructional methods.

Achievement Outcomes

Results on the various problem-solving tests did not indicate significant advantages for the PT group. The latter class began the year as the lowest of the three classes in physics aptitude and mathematics experience, and completed the year in the same position as measured by examination scores. This outcome should not be interpreted

negatively, but as inconclusive in view of the constraints of the present program context.

First, the *Tutor* was utilized for an extremely small proportion of the instructional time during the year. It thus seems unreasonable to expect it to impact performances that are likely to be influenced by so many confounding variables (e.g., student preparation for tests, learning from conventional teaching methods and other resources, homework, practice problems, cooperative learning, learner characteristics, class size, etc.).

Second, subjects consisted exclusively of honors students, thus representing a relatively homogeneous sample of students who generally perform well academically. Chances of treatment differences occurring with this select population would be greatly reduced relative to using low- or average-achievers.

Third, chance factors resulted in an uneven distribution of class sizes and abilities in the three comparison groups. The control (first-period) class, for example, was the smallest and highest in ability of the three.

Fourth, the *Tutor* was primarily designed to influence students' *approaches* to problem-solving, by reinforcing systematic solution steps and recording of step-by-step mathematical work. Such influences may increase "general" or long-term problem-solving ability, while having little impact on immediate test scores based on a restricted, known group of problems. Seemingly, short-cut or idiosyncratic solution strategies can be used on the latter problems with much greater success than would be case for transfer problems from many different domains.

Based on the above considerations, we consider the evaluation of *Tutor* effects on course achievement to be inconclusive. Further evaluation is recommended, using matched, regular physics classes over a longer treatment period.

Problem-Solving Strategy

Results from the analyses of problem-solving strategies were highly suggestive regarding the potential of the *Tutor* to influence *how* students solve problems. These results were also consistent with both student and teacher perceptions of *Tutor* effects.

The pattern of behavior indicated by the problem-solving measure revealed little change by the PT group early in the semester, in solving the beginning one- and two-dimensional kinematics problems. Consistent with their low standing on the pre-course variables, the PT group had the lowest scores of the three groups in representing relevant elements in their problem solutions. As the semester progressed, and experiences with the *Tutor* increased, this pattern reversed to reflect increasing use by the PT group of the problem elements. On the last and most advanced unit, two-dimensional dynamics, the PT students had the highest problem-solving strategy scores of the three groups.

These results suggest that the Tutor is effective at developing more systematic problem-solving strategies. Such influences, however, take time to incubate and may therefore may not emerge until later units of study. It is also noteworthy that the no-feedback control group scored relatively high on the strategy measure. The implication is that the use of print versions of the *Tutor* problems, coupled with teacher feedback, was also effective (perhaps more in the beginning, since such feedback may have implicitly or explicitly conveyed the teacher's expectancy to see complete representations used on subsequent problems.) However, the disadvantage relative to the *Tutor*, in the long-run is the difficulty for teachers of adapting feedback to individuals as problem complexity and the variability of student needs increase over time.

Recommendations

Despite the failure to show differences in problem-solving (test) scores, the evaluation findings, overall, were positive regarding the Tutor's utility for instruction and effectiveness for student learning. The following recommendations are made for further

development of and research with the *Tutor*.

1. The developers should review the student and teacher evaluations regarding the user interface, and make whatever minor improvements they consider desirable.

2. Cooperative learning should continue to be recommended as a desirable instructional strategy for using the *Tutor*.

3. A user's manual and teacher's manual (or a combination) should be developed to provide clear information about the program (instructions, set-ups, operating procedures, etc.) and guidelines for instructional uses.

4. Additional evaluation should be conducted at different schools, using regular (not honors) classes that are matched in ability and size.

REFERENCES

A. K. Bednar, D. Cunningham, T. Duffy, & J. Perry. (1991). Theory into practice: How do we link? In G. J. Anglin (Ed.), *Instructional technology: Past, present, and future* (pp. 88-101): Englewood, CO: Libraries Unlimited.

M. T. H. Chi, P. J. Feltovich, & R. Glaser (1981) Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5*, 121-152.

F. J. Di Vesta, & L. P. Rieber (1987). Characteristics of cognitive engineering: The next generation of instructional systems. *Educational Communications and Technology Journal*, 35, 213-230.

R. B. Loftin(1990). *The Intelligent Physics Tutor: Project description*. Unpublished manuscript. University of Houston-Downtown, Department of Natural Sciences.

S.. M. Ross, & G. R. Morrison (1993). Personalization of feedback in computer-based instruction. In G. Sales & J. Dempsey (Eds.). *Feedback and interactive instruction*. (pp. 187-196) Englewood Cliffs, NJ: Educational Technology Publications.